



UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

DIPARTIMENTO DI FISICA E ASTRONOMIA 'GALILEO GALILEI'

Corso di Laurea in

Fisica

**Soppressione del fondo tramite analisi di pulse
shape per rivelatori al germanio con contatti
puntiformi**

**(Background suppression through analysis of pulse
shape for germanium detectors with point-like contact)**

Relatore

Giovanni Donati

Alberto Garfagnini

Co-relatore

Katharina von Sturm

ANNO ACCADEMICO 2013/2014

Indice

1	Introduzione	3
1.1	L'esperimento GERDA	3
1.2	Rivelatori al germanio	4
1.3	Distinzione segnale-fondo	5
2	Acquisizione dati e software utilizzato	7
2.1	Acquisizione dati	7
2.2	GELATIO	8
2.3	Rete neurale e TMVA	8
3	Analisi	12
3.1	Metodologia	12
3.2	Cura dei dati e tagli applicati	16
3.2.1	A/E	17
3.2.2	Risetime	19
3.3	Tagli utilizzati	20
4	Risultati	21
4.1	Risultati allenamento	21
4.1.1	Fase di test	24
4.2	Conclusioni	26
	Bibliografia	27

Capitolo 1

Introduzione

1.1 L'esperimento GERDA

L'esperimento GERDA [1] (GERmanium Detector Array), situato nei Laboratori Nazionali del Gran Sasso (LNGS), INFN, Italia, è stato pensato e disegnato con l'intento di rivelare il doppio decadimento beta senza emissione di neutrini ($0\nu\beta\beta$) del ^{76}Ge , espresso dalla seguente relazione:

$$(A,Z) \longmapsto (A,Z+2) + 2e^{-}$$

in cui due neutroni all'interno del nucleo decadono in 2 protoni con emissione di due elettroni. Esso si distingue dal doppio decadimento beta con emissione di neutrini, per la mancanza di questi ultimi e la scoperta di un tale evento sarebbe di fondamentale importanza per la fisica particellare in quanto proverebbe la non conservazione del numero leptonico e la natura Majorana del neutrino, una particella che coincide con la sua antiparticella [2]. Una delle prime difficoltà in questo tipo di esperimento è legata al tempo di dimezzamento dei nuclei che possono produrre doppi decadimenti beta, stimato essere compreso fra i 10^{19} - 10^{24} anni, un periodo relativamente molto grande se paragonato ai tempi di sperimentazione. Per lo studio di eventi così rari risulta quindi indispensabile: (i) ridurre il più possibile gli eventi di fondo, (ii) avere grandi masse di nuclei radioattivi e (iii) effettuare misure per tempi sufficientemente lunghi. Per far fronte a queste esigenze l'esperimento GERDA è stato progettato utilizzando diodi al germanio arricchiti all' 86 % di ^{76}Ge [1], così da poter essere utilizzati allo stesso tempo sia come sorgenti che come rivelatori. Essi sono posizionati all'interno di un criostato riempito di 64 m^3 di argon liquido (LAr) atto a mantenere alle temperature di esercizio i rivelatori, e fungere da scudo a possibili radiazioni proveniente dalle pareti del criostato. Il tutto è infine inserito all'interno di un recipiente di acqua pura (590 m^3) al

fine di schermare ulteriormente l'esperimento da neutroni e raggi γ esterni e per agire infine anche come rivelatore Cherenkov per respingere eventi dovuti a muoni ad alta energia [4].

1.2 Rivelatori al germanio

In GERDA vengono utilizzati due tipi di rivelatori entrambi al germanio la cui ridotta differenza energetica tra banda di valenza e banda di conduzione (che in questo caso è 0.67eV) fa sì che essi appartengano alla classe dei semiconduttori. Come è noto, in generale i semiconduttori si differenziano in tipo p e in tipo n a seconda che i portatori maggioritari di carica siano rispettivamente le lacune o gli elettroni liberi nati dal drogaggio del materiale o dalle impurità inevitabilmente presenti. Una loro unione crea una situazione di non equilibrio in cui gli elettroni tendono a muoversi, per diffusione, dalla zona n alla zona p, e le lacune in senso opposto, lasciando entrambi dietro di sé ioni carichi che danno così vita ad un campo elettrico. Focalizzandosi in particolare nelle vicinanze della zona di giunzione, ove è molto probabile la ricombinazione elettrone-lacuna, si trova la cosiddetta regione di svuotamento, caratterizzata dall'assenza di cariche libere. E' proprio in questa regione che possono essere generate cariche al passaggio di una radiazione ionizzante. Esse migreranno poi per effetto del campo elettrico di barriera e saranno quindi raccolte originando un segnale elettrico. Tipicamente per ingrandire le aree di raccolta del segnale si applicano delle alte tensioni ai capi delle zone di svuotamento, proprio come è stato fatto in questo esperimento.

Tornando quindi ai rivelatori utilizzati in GERDA, essi sfruttano proprio la connessione fra i semiconduttori di tipo p e i semiconduttori di tipo n, differendo fra loro in particolare per la forma e quindi per il campo elettrico interno prodotto. Come illustrato in Figura 1.1 i rivelatori semi-coassiali possiedono infatti una simmetria cilindrica e presentano un foro lungo l'asse di simmetria che agisce come elettrodo p^+ , dando così luogo ad un campo elettrico radiale. Nei rivelatori BEGe (Broad Energy Germanium) l'elettrodo p^+ è posizionato sul fondo del rivelatore, disgiunto dal contatto n^+ per mezzo di un groove, ed il campo elettrico da essi ottenuto è più irregolare. Quando una radiazione interagisce con il rivelatore le coppie elettrone-lacuna originanti dai depositi di energia, cominciano a muoversi in verso opposto lungo i campi elettrici interni e sono poi raccolte dagli elettrodi, dei quali si analizza solo il contatto p^+ [5]. Un terzo rivelatore utilizzato è il P-type Point-contact Detector (PPT), un'evoluzione del BEGe da cui riprende le forme, ingrandendole, e che si distingue per il contatto molto piccolo posto sul fondo [9].

Risulta quindi evidente come la forma del segnale di carica ricavato in funzio-

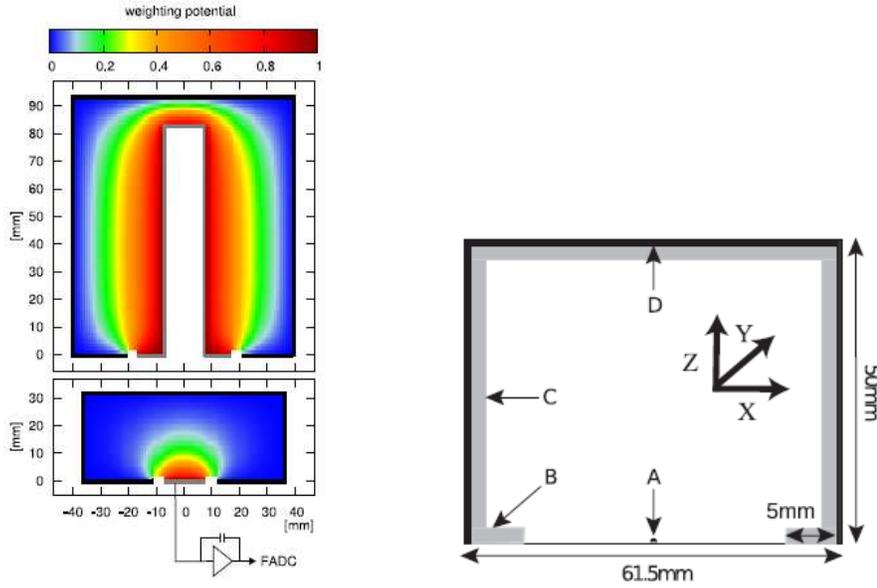


Figura 1.1: Geometrie dei rivelatori descritti: a sinistra il semi-coassiale in alto e il BEGe in basso, a destra il PPC

ne del tempo dipenda, per ogni evento, dalla distanza dagli elettrodi a cui è avvenuto il rilascio di energia, dal numero e dalla grandezza delle aree in cui esso è avvenuto, e da quanta energia è stata rilasciata. Proprio per questo motivo interazioni di natura differente vengono registrati come differenti forme d'impulso. L'obiettivo di questo lavoro sarà proprio quello di sfruttare le distinzioni fra forme di segnale registrate (pulse shape analysis) per riuscire a distinguere gli eventi che si ipotizza essere di fondo da quelli di segnale.

1.3 Distinzione segnale-fondo

Per comprendere quali eventi classificare come segnale e quali come fondo è essenziale capire che tipo di evento stiamo ricercando. Come detto il doppio decadimento beta senza emissione di neutrini può essere espresso come :

$$(A,Z) \mapsto (A,Z+2) + 2e^-$$

e si differenzia da due singoli decadimenti beta per l'assenza dei due neutrini, i quali si annichilano. Se nel singolo decadimento beta l'energia con la quale viene rilasciato l'elettrone varia a seconda dell'energia con cui viene prodotto il neutrino (per la conservazione dell'energia) così non avviene nel decadimento da noi ricercato, ove i due elettroni vengono generati con la massima energia

possibile in direzioni opposte (per la conservazione del quadrimomento). Tale energia è nota ed è data dalla differenza energetica a riposo fra i due nuclei iniziali e finali, che nello specifico vale $Q_{\beta\beta} = (2039.061 \pm 0.007)$ KeV [3], e viene assorbita dal rivelatore in pochi mm. La creazione di carica che ne consegue è quindi localizzata in una piccola regione di spazio nel rivelatore e per questo motivo si parla di Single Site Event (SSE). Gli eventi ritenuti di fondo hanno invece la caratteristica di rilasciare la loro energia in più aree del rivelatore, e si denominano per questo motivo Multiple Site Event (MSE) [5]. Si pensi ad esempio a scattering Compton multipli, in cui un fotone incidente il rivelatore interagisce con un elettrone cedendo parte della sua energia, e procede al suo interno riuscendo ad urtare altri elettroni e cedendo loro ulteriore energia rimanente. Immaginando il tutto si intuisce come l'andamento della carica raccolta vari nel tempo nei due casi in modo molto differente. Dall'analisi dei segnali

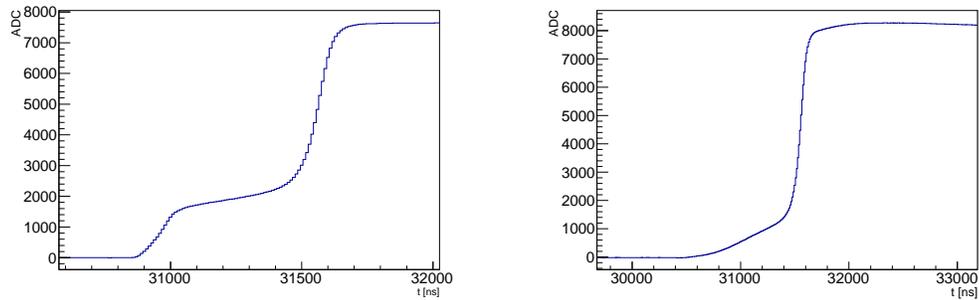


Figura 1.2: Esempi di interazione nel rivelatore BEGe: A sinistra un esempio di MSE e a destra un esempio di SSE

MSE e SSE è possibile trovare dei parametri tramite i quali, grazie all'utilizzo di una rete neurale, sarà possibile classificare eventi ignoti nei due tipi di eventi descritti.

Capitolo 2

Acquisizione dati e software utilizzato

2.1 Acquisizione dati

In GERDA i raw data vengono campionati da flash-ADCs (FADC) a 14 bit funzionanti ad una frequenza di campionamento di 100 MHz. Per ogni evento fisico di trigger i segnali presenti in tutti i rivelatori vengono acquisiti così da poter ricercare eventuali coincidenze. I dati acquisiti durante l'esperimento sono organizzati in una struttura a più livelli, come illustrato in Figura 2.1, ognuno dei quali chiamato tier seguito dal numero corrispondente il livello:

I raw data occupano il livello più basso (tier 0). Essi vengono innanzitutto

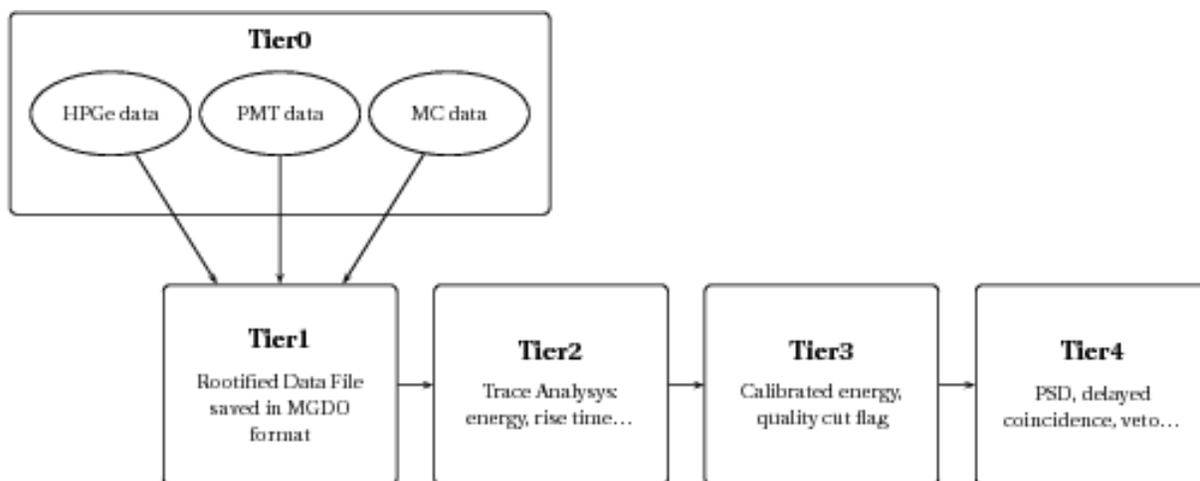


Figura 2.1: Struttura a Tier dei dati GERDA

convertiti in un nuovo formato e memorizzati come Tier1, per poi essere distribuiti e analizzati. Tier0 e Tier1 contengono esattamente le stesse informazioni, con la differenza che il Tier1 risulta essere un root-file, formato standardizzato progettato per essere più flessibile, maneggevole e facilmente leggibile. In esso si ritrovano esattamente le tracce di ogni segnale rilevato, o in altre parole l'andamento della carica rilasciata nel rivelatore in funzione del tempo. Si giunge quindi al Tier2 applicando al Tier1, e quindi a tutte le tracce in esso salvate, processi digitali in grado di prelevare le informazioni di interesse, come per esempio il tempo di salita del segnale, la sua ampiezza ecc. . . . Similmente si crea il Tier3 estraendo e processando informazioni dal Tier2 e così arrivando a Tier di livello via via più alto [6]. In particolare presteremo un attimo di interesse al passaggio Tier1-Tier2, reso possibile grazie ad un software chiamato GELATIO.

2.2 GELATIO

GELATIO [6] è un ambiente di analisi appositamente disegnato per fornire una completa suite di strumenti atti a estrarre informazioni dai Tier1 e memorizzarle nei Tier2. Esso presenta un'interfaccia dall'approccio modulare: ogni modulo è associato ad uno specifico algoritmo e permette di estrarre dalle tracce dei segnali del Tier1 una determinata informazione. Si ha così la possibilità di scegliere in sequenza i moduli di interesse e applicarli al Tier1 per ottenere il Tier2 con le grandezze che si desiderano. Saranno precisamente alcune di queste grandezze che nel nostro caso verranno fornite come parametri di input alla rete neurale per discriminare tra eventi di segnale e background. La lista dei moduli attivati e dei parametri utilizzati vengono quindi salvati dall'utente in un apposito file (file.INI).

2.3 Rete neurale e TMVA

TMVA (toolkit for multivariate data analysis) [7] è un software integrato in root ed è caratterizzato da un ambiente di apprendimento composto di diversi classificatori di analisi multivariata. Ognuno di essi usa la programmazione orientata ad oggetti C++/ROOT e appartiene alla famiglia degli algoritmi di apprendimento supervisionato: in primo luogo vengono allenati su un set di dati il cui risultato di output è noto, definendo un valore di taglio, per poi essere applicati e classificare dati ignoti sulla base di questo valore. Fra essi figurano le reti neurali artificiali (ANN : artificial neural networks) di cui sono disponibili tre implementazioni. In questo lavoro verrà utilizzata la Multi

Layer Perceptron (MLP), consigliata dalla stessa TMVA perchè di più recente sviluppo .

In generale una rete neurale è un modello computazionale ispirato all'architettura e ai principi di funzionamento del sistema nervoso centrale umano: possiede cioè una sua intelligenza artificiale che le conferisce le abilità di imparare e formulare giudizi. Essa è composta di tante unità computazionali, rappresentanti i neuroni, interconnessi fra loro con un *peso* che si modifica ogni qualvolta un'informazione di input viene captata dalla rete. Come si os-

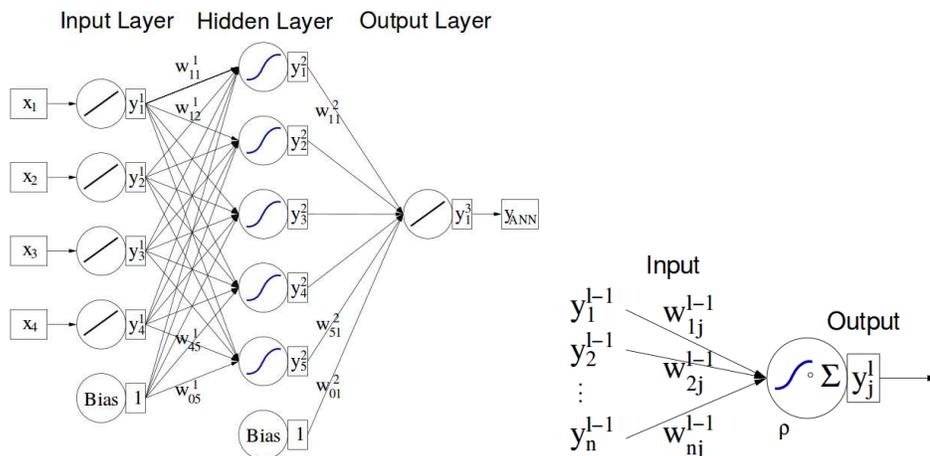


Figura 2.2: A sinistra : architettura di un Multi Layer Perceptron. A destra : un singolo neurone j linkato da vari neuroni appartenenti al layer precedente

serva dalla Figura 2.2 la peculiarità della rete MLP utilizzata in questo lavoro risiede nel fatto che i neuroni sono organizzati in più layers e che le connessioni permesse avvengono solo fra neuroni di layers differenti. Il primo layer è l'input layer, l'ultimo è l'output layer, e tutti i restanti sono hidden layers. In un problema di classificazione con n variabili si costruisce una rete con lo stesso numero di neuroni nello strato di input, e un solo neurone di output. La Figura 2.2, a destra, mostra in particolare il funzionamento di un generico neurone j appartenente all' l -esimo layer: esso è connesso ad ogni neurone i del layer precedente ($l - 1$), ricevendo da ciascuno un input Y_i^{l-1} . L'importanza della connessione è raccolta dalla variabile peso w_{ij}^{l-1} . Ogni neurone è inoltre associato ad un bias value, w_{0j}^l , che indica la sua soglia di attivazione. Il modo in cui ciascun neurone processa le informazioni di input e fornisce un certo output è descritto dalla *neuron response function* (ρ), che può essere vista come una funzione $\rho : \mathbb{R}^n \mapsto \mathbb{R}$ che mappa gli n neuroni di input in un neurone di output. Essa può essere più o meno complessa a seconda del problema da risolvere, ed è composta di due funzioni: la *synaps function* $k : \mathbb{R}^n \mapsto \mathbb{R}$, e

la *neuron activation function* $\alpha : \mathbb{R} \mapsto \mathbb{R}$, così che $\rho = \alpha \circ k$. Quelle utilizzate dalla nostra rete sono:

$$k : (y_1^{(l)}, \dots, y_n^{(l)} | w_{0j}^{(l)}, \dots, w_{nj}^{(l)}) \mapsto w_{0j}^l + \sum_{i=1}^n y_i^{(l)} w_{ij}^{(l)} \quad (2.1)$$

$$\alpha : x \mapsto \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.2)$$

Prima di essere applicate le reti neurali necessitano di una fase di allenamento, atta a settare nel migliore dei modi i pesi e le connessioni fra i neuroni. Un algoritmo, detto di *backpropagation*, si occupa di tutto ciò. Sostanzialmente un set di N eventi noti viene fornito alla rete come input, e per ciascuno di essi (sia a l'evento a -esimo) la rete fornisce un output $Y_{ANN,a}$, che viene quindi confrontato col valore atteso noto y_a : 0 per background e 1 per segnale. I pesi vengono quindi scelti in modo da minimizzare la seguente funzione E (*error function*):

$$E(\vec{x}_1, \dots, \vec{x}_N | \vec{w}) = \sum_{a=1}^N \frac{1}{2} (y_{ANN,a} - y_a)^2, \quad (2.3)$$

ove \vec{x}_i rappresentano i valori assunti dalle variabili di discriminazione durante l'evento a , e \vec{w} denota l'insieme di tutti i pesi assunti dalla rete in quel momento. Il set di pesi che minimizza la error function E viene trovato utilizzando il metodo di *steepest or gradient descent*, che assume per ipotesi la differenziabilità della response neuron function rispetto ai pesi di input.

In questo lavoro ciò che si vuole ottenere è una discriminazione tra eventi di segnale (SSE) ed eventi di fondo (MSE). Se prendendo una sola variabile di discriminazione è facile settare un taglio che distingua il segnale dal fondo, così non è utilizzandone, come nel nostro caso, più di una. La potenza di una rete neurale allenata entra in gioco proprio qui: essa trasforma un problema multidimensionale a più variabili in un problema ad una sola variabile, associando ad ogni evento un certo valore compreso tra 0 e 1. Diventa così facile settare un valore di taglio idoneo, al di sotto del quale classificare un dato evento come background e altrimenti come segnale. La Figura 2.3 rappresenta la tipica risposta di una ANN: ogni valore associato dalla rete agli eventi di training vengono utilizzati per costruire i due istogrammi riportati. Si intuisce che più i picchi sono separati e vicini a 0 e 1 e meglio la rete riesce a svolgere il suo ruolo di discriminazione.

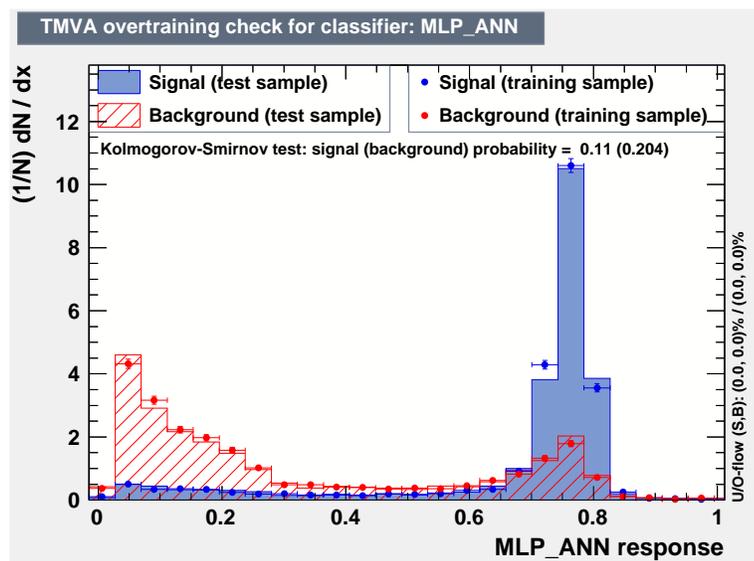


Figura 2.3: Risultati allenamento di una ANN. Se la rete funziona bene l'istogramma del background sarà piccolo vicino a 0 e l'istogramma del segnale vicino a 1

Capitolo 3

Analisi

3.1 Metodologia

L'obiettivo di questo lavoro è riuscire a discriminare eventi di segnale da eventi di fondo utilizzando una rete neurale, per due diverse tipologie di rivelatori: i BEGe e i PPC. Ciò richiede anzitutto avere un set di dati già noto da poter utilizzare come campione di allenamento per la rete e a tal fine sono stati utilizzati i dati ottenuti esponendo i rivelatori ad una sorgente nota di torio (^{228}Th). Come già citato classifichiamo eventi di segnale ed eventi di fondo a seconda che siano rispettivamente dei SSE (Single Site Event), come il doppio decadimento beta ricercato, o dei MSE (Multiple Site Event). L'appartenenza di un evento ad una delle due tipologie è strettamente legata al tipo di interazione che ha dato lui luogo all'interno del rivelatore. Dalla teoria si conoscono essenzialmente 3 modi in cui la radiazione può interagire con la materia, dipendenti dalla sua energia: effetto fotoelettrico, effetto Compton e produzione di coppia. La Figura 3.1 mostra le possibili interazioni dei fotoni con la materia. Nell'effetto fotoelettrico l'energia del fotone viene completamente assorbita dall'elettrone con cui urta per la sua prima ed unica volta, ed è quindi rilasciata localmente in un certo punto (SSE). Al crescere dell'energia del fotone aumenta la probabilità che esso produca scattering Compton: qui il fotone incidente può in generale attraversare il rivelatore interagendo con uno (SSE) o più elettroni (MSE) cedendo ciascuno parte della sua energia, per poi uscirne o essere completamente assorbito tramite effetto fotoelettrico. L'energia viene così rilasciata in tutti quei siti in cui è avvenuto l'urto. Infine se l'energia del fotone è maggiore di 1022 KeV può verificarsi il fenomeno della produzione di coppia, un processo in cui il fotone si trasforma in una coppia elettrone-positrone di cui il primo viene velocemente assorbito (SSE) e il secondo, annichilendosi, produce una nuova coppia di fotoni da 511 KeV. Si parla dunque di double escape peak (DEP)

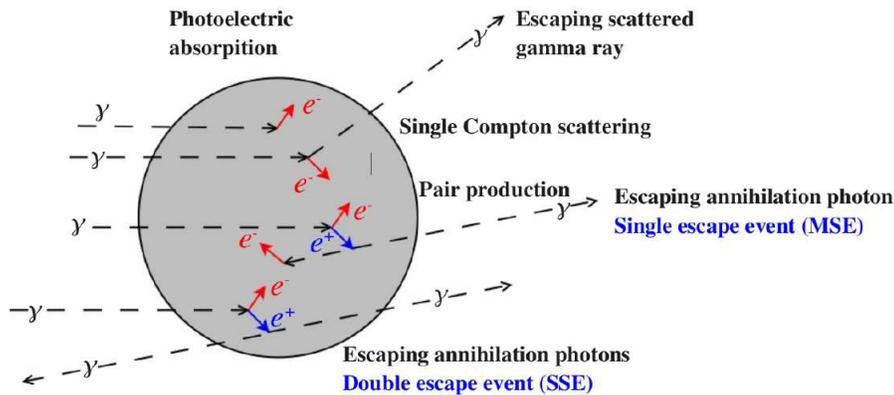


Figura 3.1: interazioni possibili tra fotone e materia: effetto fotoelettrico, Scattering Compton e produzione di coppia

quando entrambi i fotoni riescono ad uscire dal rivelatore inalterati senza interagire (SSE), e di single escape peak (SEP) nel caso in cui solamente uno vi riesca (MSE).

E' possibile individuare facilmente tali tipi di interazioni nello spettro del ^{228}Th , visibile in Figura 3.2, osservandone in particolare i seguenti picchi partendo da destra:

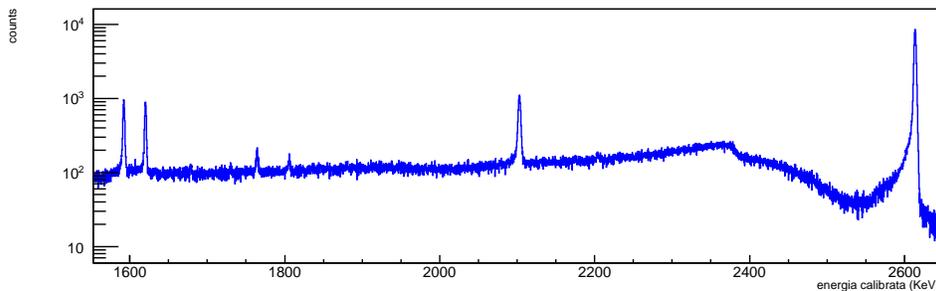


Figura 3.2: Spettro calibrato della sorgente di ^{228}Th acquisito con un PPC

- 2614 KeV - Full energy peak (FEP): tutta l'energia del fotone emesso dal ^{228}Th viene rilasciata nel rivelatore, tramite in generale una combinazione delle interazioni sopra descritte.
- 2103 KeV - Single escape peak (SEP -): a seguito della produzione di coppia, uno dei due fotoni dovuti all'annichilazione del positrone riesce

ad evadere il rivelatore. L'energia del picco è quindi 511 KeV minore del FEP.

- 1620 KeV - Picco dovuto al ^{212}Bi . Si tratta in particolare di un full energy peak.
- 1592 KeV - Double escape peak (DEP): entrambi i fotoni dovuti all'annichilazione del positrone evadono il rivelatore indisturbati. L'energia del picco è quindi 1022 KeV minore del FEP del ^{228}Th .

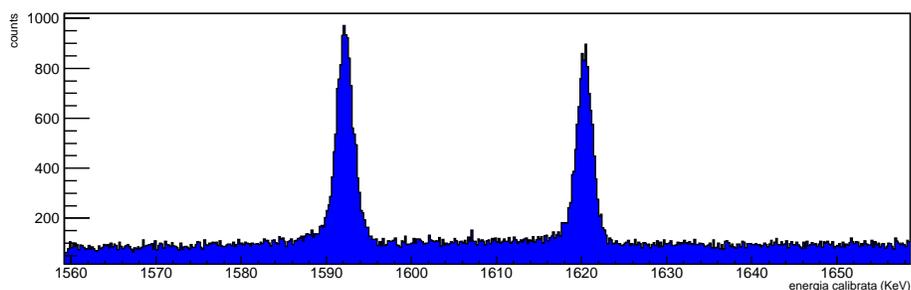


Figura 3.3: I due picchi utilizzati per identificare i dati che serviranno per allenare la rete: a sinistra il DEP e a destra il picco del bismuto

Si è scelto di utilizzare come dati di segnale da fornire alla rete per allenarsi i SSE provenienti dal DEP (1592 KeV) e come dati di background quelli presenti invece nel picco del ^{212}Bi (1620 KeV), composto quasi esclusivamente da MSE (figura 3.3).

Operativamente si è eseguito un fit gaussiano di entrambi i picchi selezionati e acquisiti tutti i dati all'interno di una finestra di 2σ dal centroide. Utilizzando quindi l'apposito software GELATIO si sono estratte, per ognuno di questi eventi, le grandezze utili alla nostra analisi, memorizzandole in un root-file chiamato Tier2.

Il passo successivo è stato scegliere quali parametri utilizzare come variabili di discriminazione per la rete neurale. In questo lavoro ne sono state scelte 3:

- A/E : rapporto tra ampiezza della derivata del segnale di carica, ovvero il segnale di corrente, e la sua energia.
- risetime : tempo che il segnale di carica impiega a salire dal 10% al 90% della sua altezza massima.

- Un parametro caratterizzante l'asimmetria del segnale di corrente, definito nel seguente modo:

$$Asym = (A/E) \cdot \frac{L - R}{L + R} \quad (3.1)$$

ove L e R vengono calcolate sul segnale di corrente e rappresentano rispettivamente l'integrale dalla parte sinistra e destra dell'ampiezza massima e spigoli inferiori il 3% dell'ampiezza massima.

La scelta risulta dettata dal fatto che si sono scelti i picchi SSE e MSE molto vicini energeticamente: gli eventi che vi appartengono sono allora caratterizzati dall'aver generato nel rivelatore un numero di portatori di carica simile. Essendo quest'ultima rilasciata in più aree del rivelatore nei MSE, sarà allora raccolta non contemporaneamente, come accade nei SSE, ma in più step temporali. Le differenze si rifletteranno allora nei segnali di corrente i quali, a parità di energia, avranno più di un picco e ampiezza minore nei MSE, e un solo picco di ampiezza maggiore nei SSE. Si intuisce come un ragionamento del tutto simile possa essere fatto per i parametri risetime e asimmetria [8].

Terminata questa fase si è proseguito individuando e settando la rete neurale che sarà poi allenata e applicata. Fra le tre candidate implementate in TMVA, come già scritto, si è optato per utilizzare la Multi Layer Perceptron (MLP) organizzandola in tre strati, come mostrato in Figura 3.4: il primo, quello di input, contiene un numero di neuroni pari al numero delle variabili di discriminizzazione scelte più un neurone ulteriore di bias; il secondo strato, è composto da 8 neuroni; il terzo strato, è costituito da un unico neurone di output. Ci si aspetta che una volta allenata la rete, quest'ultimo, dovendo fornire una risposta continua di valori compresi tra 0 e 1, restituisca valori prossimi a 0 per i dati di background proveniente dal picco del bismuto e valori prossimi a 1 per i dati di segnale provenienti dal DEP.

A questo punto la rete allenata è in grado di riconoscere e giudicare un segnale qualsiasi basandosi sulle stesse tre variabili utilizzate per il suo allenamento e attribuisce lui un valore compreso tra 0 e 1 a seconda che lo reputi più affine ad un evento di background o di segnale. Si è quindi definito un valore di taglio, detto *cut value*: ogni volta che la rete giudicherà un evento con un valore minore del cut value, esso verrà etichettato background, altrimenti segnale. La scelta del cut value risulta a questo livello di analisi arbitraria. Infatti, se da un lato si è pensato agli eventi appartenenti al DEP e al picco del bismuto per allenare la rete, basandosi sul tipo di interazioni che possono incontrare i fotoni emessi dalla sorgente di ^{228}Th , dall'altro non sappiamo quanti effettivamente

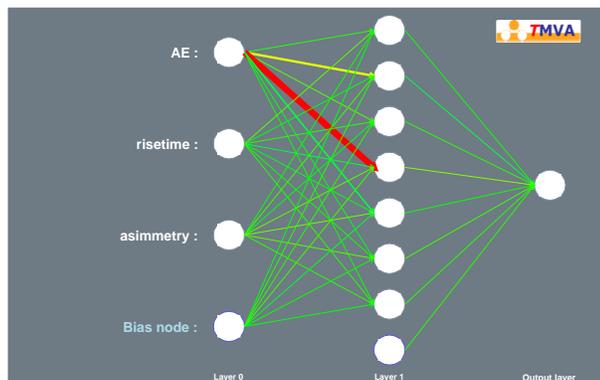


Figura 3.4: Architettura delle rete neurale MLP utilizzata. Sono presenti 3 neuroni di input, uno strato intermedio avente 8 neuroni, e un neurone di output. Il colore e la larghezza di ogni freccia indica il valore e il peso delle connessioni.

siano conseguenza del decadimento del ^{228}Th e quanti no. Ciò richiederebbe simulazioni montecarlo eseguite a priori tenenti conto delle altre radiazioni a cui possono essere esposti i rivelatori.

Determinato il cut value si valuta quindi l'efficienza della rete andando a calcolare quanti eventi vengono giustamente classificati come segnale e background: si contano cioè gli eventi del DEP che vengono riconosciuti come segnale (*signal efficiency*) e come fondo (*signal rejection*) e quelli del picco del bismuto che sono classificati come segnale (*background rejection*) e come fondo (*background efficiency*).

Terminata la fase di allenamento e scelto un cut value idoneo, è infine necessario testare la rete per capire se è davvero in grado di discriminare un campione di dati sconosciuto. A tal scopo la rete allenata è stata applicata sugli eventi appartenenti al SEP, ritenuti per la maggior parte dei multiple site event: ci si aspetta quindi che la maggior parte venga riconosciuta come background.

3.2 Cura dei dati e tagli applicati

Delle tre variabili di discriminazione scelte il rapporto A/E e la risetime necessitano di una certa attenzione, atta a migliorare le prestazioni della rete neurale in fase di allenamento e applicazione ad un campione di dati ignoto.

3.2.1 A/E

La variabile A/E ha come denominatore l'energia dell'evento rilevato. I dati scelti per allenare la rete appartengono a due picchi che sono energeticamente simili ma non uguali, e questo fa sì che se A/E presenta una dipendenza dall'energia allora la rete possa incorporare tale differenza nel modo di discriminare il segnale dal fondo. Applicata poi ad un campione di dati con energie che variano in un range assai più ampio rispetto a quello utilizzato per allenarsi, si rischia che i risultati possano essere fuorvianti per via di questo effetto. Quello che si è fatto è stato allora verificare la dipendenza del rapporto A/E dall'energia per entrambi i rivelatori, e in caso positivo apportare una cura ai dati che la elidesse. In figura 3.5, figura 3.6 e figura 3.7 sono rappresentati istogrammi bidimensionali in cui nell'asse delle ascisse si è posto l'energia calibrata e in quello delle ordinate il rapporto A/E. Si nota subito come ci siano

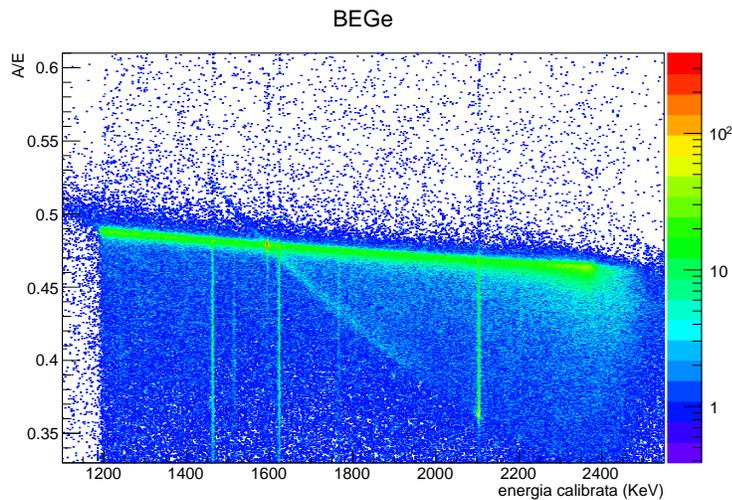


Figura 3.5: In verde i SSE: si nota come l'andamento di A/E per i SSE dipenda dall'energia.

zone in cui le combinazioni delle due variabili sono più frequenti (in verde): esse rappresentano i SSE. Immaginando infatti di suddividere l'energia in tanti intervalli si ha che per ognuno di essi gli eventi SSE che vi appartengono rilasciano nel rivelatore la stessa energia generando simili segnali di corrente e forniscono perciò rapporti A/E molto vicini fra loro. Gli eventi MSE creano segnali di corrente che possono invece avere andamenti molto diversi, e così anche i corrispondenti rapporti A/E. I SSE saranno quindi rappresentati dai bin con conteggi più alti.

Il fit è stato eseguito prendendo come valori di A/E quelli appartenenti alle

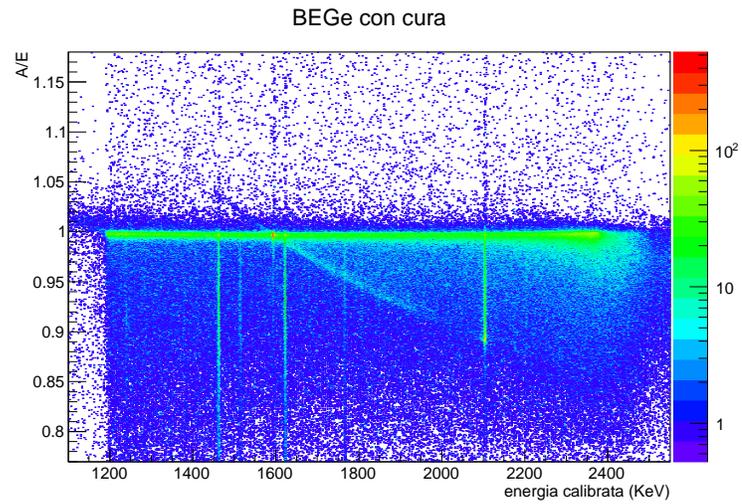


Figura 3.6: Dopo la cura l'andamento del rapporto A/E in funzione dell'energia è piatto

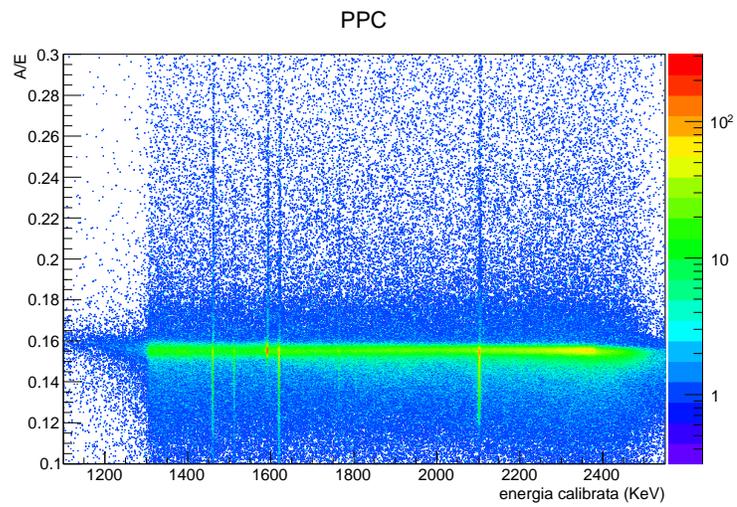


Figura 3.7: In questo caso la dipendenza ricercata non è presente e non è necessario eseguire la cura dei dati.

zone in verde e, ricavando coefficiente angolare e intercetta, si è infine potuto portare a termine la cura.

3.2.2 Risetime

Particolare attenzione è stata posta anche per la variabile di discriminazione risetime. Controllando le risetime per gli eventi appartenenti al DEP e al picco del bismuto ci si è accorti che la quasi totalità degli eventi presentavano un valore compreso tra 0 e 2000 ns, in entrambi i rivelatori, e che lo 0.1% presentava invece valori maggiori raggiungendo per esempio i 95000 ns. Si sono quindi individuate le *pulse shape* con valori di risetime anomali e analizzate. Si fornisce qualche esempio in Figura 3.8 e Figura 3.9, dove è evidente riconoscere eventi di pile-up, che si desidera scartare. Si è optato quindi per allenare la rete con i dati aventi risetime minore di 2000 ns.

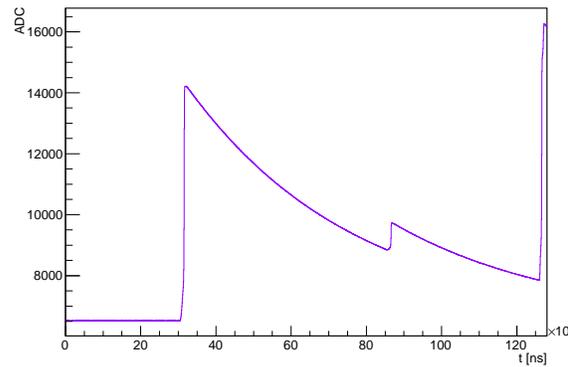


Figura 3.8: Esempio di traccia in cui la risetime ha un valore di 95680 ns. Gli eventi rilevati in questo caso sono tre rispettivamente ai tempi $30\mu s$, $90\mu s$, $125\mu s$. E' facile notare il perchè di un valore così alto: il 90 % dell'altezza massima viene raggiunto dal terzo evento raccolto, mentre il 10% già nel primo. La risetime registrata prende in considerazione quindi tre eventi e non uno.

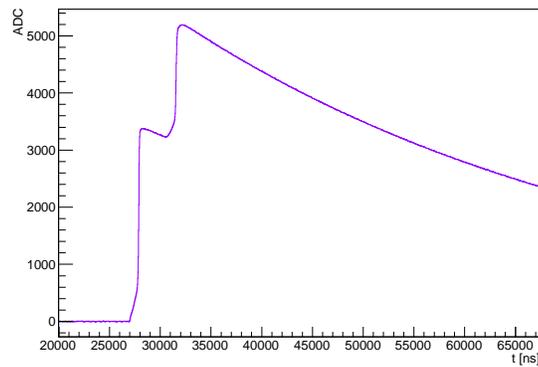


Figura 3.9: Risetime = 3920 ns. Sono presenti due eventi: raccolta la carica generata dal primo il condensatore inizia a scaricarsi ma viene bruscamente interrotto e costretto a riprendere la carica a causa degli elettroni provenienti dal secondo evento. L'altezza massima raggiunta dal segnale è quindi dovuta alla somma dei due eventi e non ad uno solo e la risetime risultante non è quella voluta.

3.3 Tagli utilizzati

Oltre al taglio sulla risetime descritto sopra sono stati utilizzati altri tre tagli, sfruttando i seguenti moduli di controllo precaricati nel software GELATIO:

- GEMDFADC - waveformTag : controlla che non siano presenti overflow di carica nelle tracce registrate.
- GEMDTrigger - Trigger : controlla che non siano presenti più di un evento in ogni traccia registrata.
- GEMDQuality - isGood : indica se sono presenti dei moduli che hanno commesso errori analizzando una certa traccia.

Capitolo 4

Risultati

4.1 Risultati allenamento

Tutte le informazioni rilevanti inerenti i risultati dell'allenamento della rete neurale sono contenute nei grafici di *classifier output distribution* in Figura 4.1:

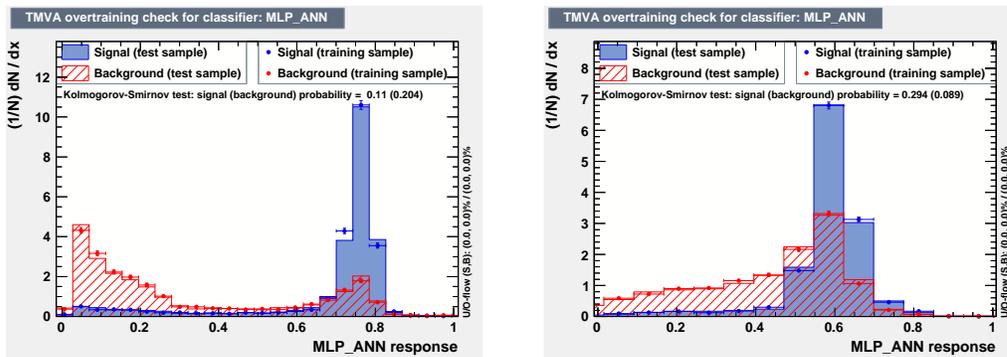


Figura 4.1: A sinistra: *classifier output distribution* per i BEGe. A destra: *classifier output distribution* per i PPC

E' immediato osservare come la rete riesca a distinguere in maniera molto più chiara il segnale dal fondo quando allenata sui dati BEGe rispetto ai PPC ove la distinzione è poco marcata.

La scelta ottimale del cut value è stata scelta in maniera da massimizzare la *signal efficiency* e la *background rejection*, definite nella sezione analisi. I risultati ottenuti sono indicati in Tabella 4.1: essi mostrano chiaramente come a parità di *signal efficiency* la *background rejection* sia minore nel rivelatore PPC, fatto già suggerito dagli istogrammi in Figura 4.1. Si è cercato di capirne il motivo andando ad analizzare le differenze fra i valori normalizzati assunti

	BEGe	PPC
signal efficiency	0.79	0.78
signal rejection	0.21	0.22
background efficiency	0.18	0.35
background rejection	0.82	0.65
cut value	0.70	0.55

Tabella 4.1: Risultati allenamento ANN per i BEGe e i PPC

dalle variabili di discriminazione per i dati di segnale e fondo per entrambi i rivelatori, riportati in Figura 4.2 e Figura 4.3.

La Figura 4.3 suggerisce in effetti che nei PPC le variabili di discriminazione

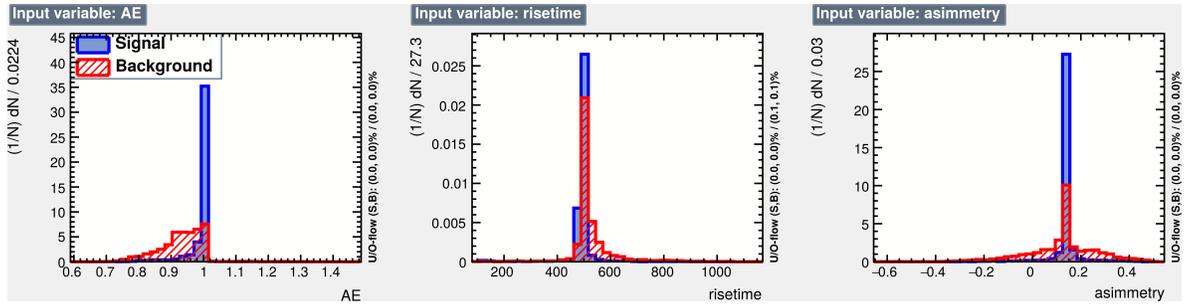


Figura 4.2: input variables per il BEGe

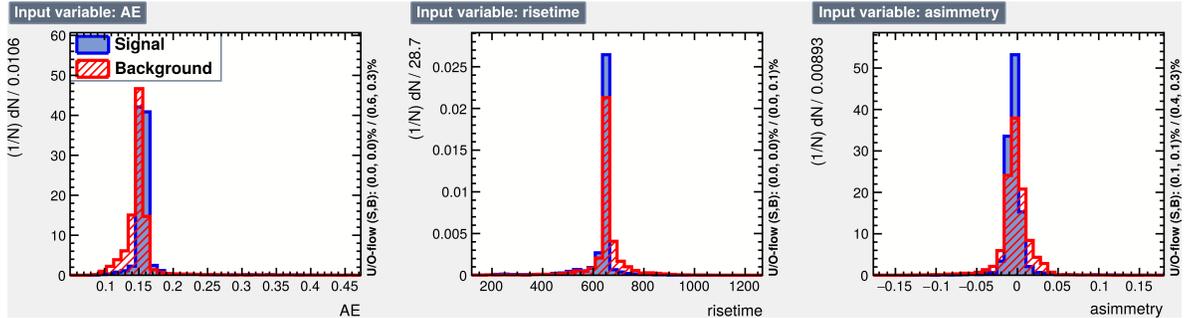


Figura 4.3: input variables per il PPC

scelte assumono valori molto simili nei campioni di segnale e fondo. Per tentare di aiutare la rete a distinguerli più egregiamente si sono aggiunte altre 50 variabili relazionate alla risetime, seguendo la seguente idea: in Figura 4.4 si riporta uno zoom dell'istogramma relativo alla risetime per il PPC, già presente in Figura 4.3. Si nota che i valori maggiori di risetime sono

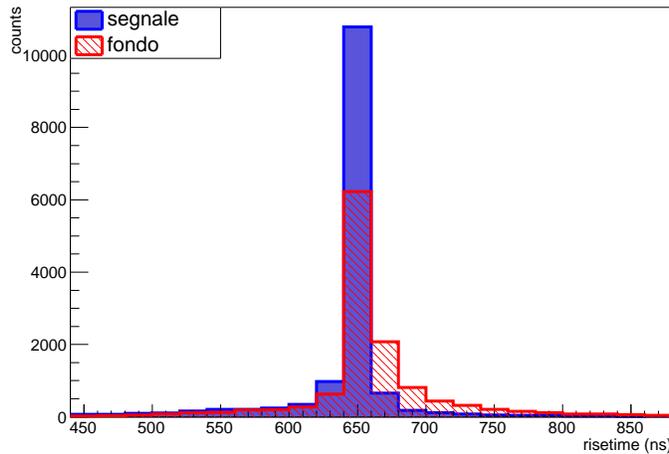


Figura 4.4: istogramma delle risetime assunte dal campione di fondo e segnale nel PPC

	BEGe	PPC
signal efficiency	0.79	0.80
background rejection	0.82	0.65
cut value	0.70	0.55

Tabella 4.2: Risultati con l'aggiunta del modulo multi-risetime

assunti con più frequenza dai dati provenienti dal campione di fondo, come ci si aspetta. Si può supporre che eventi di segnale e fondo, oltre a distinguersi per il tempo di salita dal 10% al 90% dell'altezza massima del segnale di carica, possano diversificarsi anche per risetime aventi altri estremi dell'altezza del segnale. Per esempio: 1% - 50%, 3% - 50%..., 50% - 51%, ..., 50% - 99%. Associare ad ogni step una variabile ha permesso di introdurre nell'allenamento della rete neurale 50 nuove variabili di discriminazione, che per essere ricavate per ogni evento hanno richiesto l'introduzione di un ulteriore modulo, che denomineremo qui di *multi-risetime*. La stessa aggiunta è stata fatta anche per i rivelatori BEGe.

I risultati sono riportati in Tabella 4.2: le variabili aggiunte non sembrano influenzare la capacità della rete nel discriminare gli eventi rilevati dal BEGe, il cui unico risultato è aumentare del 2% la signal efficiency nel caso del PPC. In quest'ultimo caso aver introdotto le nuove variabili porta però con sé un altro effetto: osservando la classifier output distribution fornita ora da TMVA in Figura 4.5 si nota ora che i due picchi di segnale e fondo sono più separati.

	BEGe	PPC
% eventi di fondo	0.89	0.67
% eventi di segnale	0.11	0.33

Tabella 4.3: Tagli effettuati sul SEP

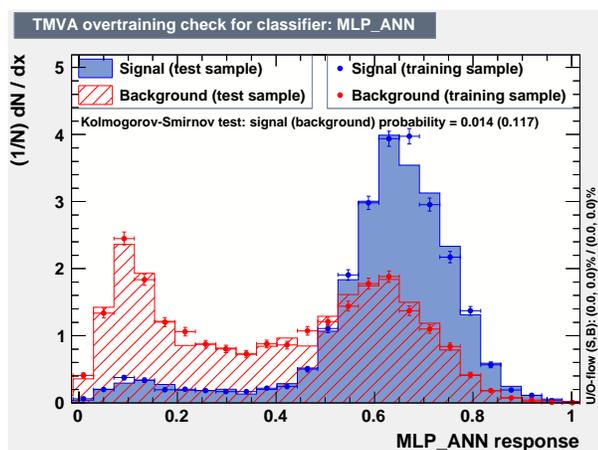


Figura 4.5: input variables per il PPC dopo l'aggiunta del modulo multi-risetime

4.1.1 Fase di test

In questa fase si è testata la rete allenata su altri dati dello spettro del torio. Si sono scelti per esempio quelli appartenenti al SEP, che a priori si ipotizzano essere dei MSE. Ci si aspetta quindi che gran parte di essi vengano riconosciuti dalla rete come fondo e vengano quindi eliminati. I risultati sono riportati in Tabella 4.3. Analoghe sono le aspettative per gli eventi appartenenti alla spalla compton dello spettro. Gli output di tali test sono riportati in figura 4.6, figura 4.7 e figura 4.8, che mostrano visivamente le differenze fra gli eventi che sopravvivono al taglio apportato dalla rete.

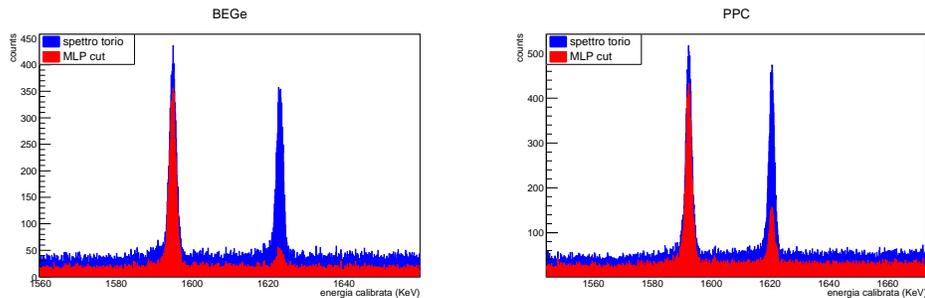


Figura 4.6: Tagli sui DEP e FEP del ^{228}Th , utilizzati come campioni di allenamento

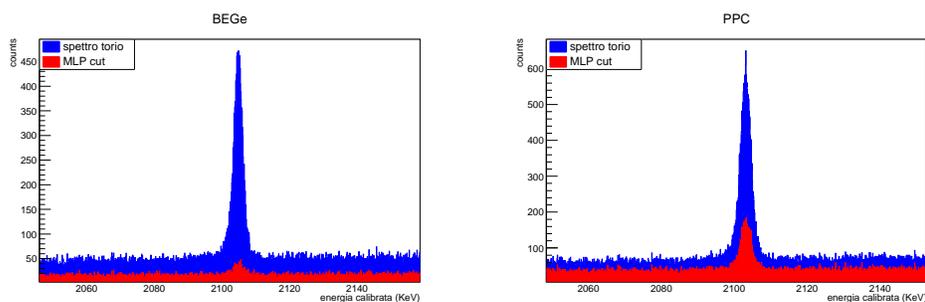


Figura 4.7: Tagli sul SEP

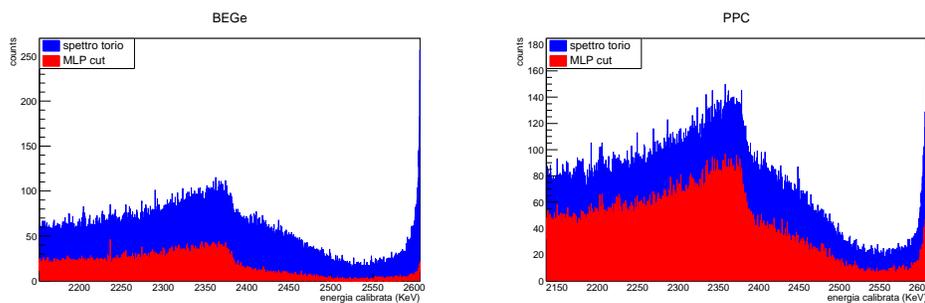


Figura 4.8: Tagli sulla spalla compton

4.2 Conclusioni

L'obiettivo del lavoro di tesi è stato riuscire a discriminare gli eventi di segnale dagli eventi di fondo nell'esperimento GERDA, considerando segnale gli eventi SSE, come il doppio decadimento beta ricercato, e come fondo gli eventi MSE. Le reti neurali si sono rivelate a tal scopo uno valido strumento: le percentuali di eventi di segnale sopravvissuti sono state del 79% e del 80% rispettivamente nei BEGe e nei PPC, e le percentuali di eventi di background eliminati del 82% e del 65% rispettivamente. Si può concludere che per i BEGe i valori sono stati in buonissimo accordo con quelli attesi [10]. Per tentare di aumentare la minor background rejection dei PPC si sono aggiunte nuove variabili di discriminazione correlate tutte ai tempi di salita dei segnali, così da aiutare la rete a trovare più elementi utili al suo lavoro. Ciò non ha portato però ad alcuna differenza nella background rejection. Si ipotizza che il motivo della differenza nei risultati ottenuti per i due rivelatori dipenda dalla loro diversa geometria di costruzione e dal loro differente campo elettrico interno. Nel caso dei PPC ciò sembra ripercuotersi in particolar modo nelle risetime degli eventi che acquisisce, che faticano a distinguersi per eventi MSE ed eventi SSE. Una possibilità per aumentare la background efficiency per i PPC potrebbe essere trovare una nuova variabile di discriminazione, utilizzando però sempre le stesse grandezze ricavate tramite GELATIO.

Bibliografia

- [1] K. H. Ackermann *et al.* (GERDA Collaboration), *Eur. Phys. J. C* 73, 2230 (2013).
- [2] S. E. Hemmer, *Study of lepton Number Conserving and Non-Conserving Processes using GERDA Phase I data*, Tesi di Dottorato, xxvi ciclo, Università di Padova.
- [3] B. J. Mount, M. Redshaw, E. G. Myers, *Double- β -decay Q values of ^{74}Se and ^{76}Ge* , *Phys.Rev.C*81, 032501, 2010.
- [4] R. Brugnera, A. Garfagnini, *Status of the Gerda Experiment at the Laboratori Nazionali del Gran Sasso*, Hindawi Publishing Corporation *Advances in High Energy Physics*, Volume 2013.
- [5] M. Agostini *at al.*, *Pulse shape discrimination for GERDA Phase I data*, *Eur. Phys. J. C* (2013) 73:2583.
- [6] M. Agostini, L. Pandola, P. Zavarise, O. Volynets, *GELATIO: a general framework for modular digital analysis of high-purity Ge detector signals*, *JINST* 6 P08013 (2011), arXiv/1106.1780.
- [7] TMVA , Toolkit for Multivariate Data Analysis with ROOT, <http://tmva.sourceforge.net>.
- [8] M. Agostini *et al.*, *Signal modeling of high-purity Ge detectors with a small read-out electrode and application to neutrinoless double beta decay search in ^{76}Ge* , *JINST* P03005.6 (2011).
- [9] E. Aguayo *at al.*, *Characteristics of signals originating near the lithium-diffused N+ contact of high purity germanium p-type point contact detectors*.
- [10] M. Perini, *Germanium detectors pulse shape analysis using artificial neural networks*, LT thesis, unipd 2013.